

# DIFFICULTIES WITH OBJECTIVE EVALUATION OF INNOVATION IN TEACHING

BU-256-M

N. S. Urquhart and W. H. Swallow

June, 1968

## ABSTRACT

Innovation in teaching manifests itself either in the development of new instructional approaches to achieve established educational goals, or in the creation of new educational goals. The two situations pose different problems. It is relevant to attempt objective comparison of teaching procedures which are alternate routes to the same goals, but such comparisons are apparently confounded with other effects. For this reason we discuss the statistical concept of confounding and illustrate its relevance within the present context. Two comparisons are suggested which, although confounding teaching methods with teacher attitudes, are meaningful. When one is revising his goals and correspondingly altering his teaching practices, subjective evaluation of goals, not comparison of teaching practices, is relevant.

---

Biometrics Unit, Department of Plant Breeding and Biometry, Cornell University, Ithaca, New York.

# DIFFICULTIES WITH OBJECTIVE EVALUATION OF INNOVATION IN TEACHING

BU-256-M

N. S. Urquhart and W. H. Swallow

June, 1968

## INTRODUCTION

When a new approach to teaching is suggested, the question often arises: "How might the effectiveness of this approach be evaluated?" Most often the questioner has in mind obtaining some objective, statistically satisfying comparison between two or more alternative procedures (perhaps comparing the proposed innovative procedure with the procedure which is currently used). This paper suggests first, that such a comparison is only relevant when the two procedures are designed to achieve the same goal, and second, that any comparison of teaching methods will necessarily be confounded with other effects. Two cases are cited where a meaningful comparison of teaching procedures (confounded with teacher attitudes) is relevant and attainable.

## CONFOUNDING IN EVALUATION

The effects of two factors are confounded when it is possible to measure only their joint impact. No amount of statistical manipulation can separate this measure of their joint impact into component parts representing estimates of the effects of the individual factors. For example, suppose two alternative teaching methods are to be implemented---method A by teacher A and method B by teacher B---and subsequently compared on the basis of student performance. Suppose the students were randomly assigned to teachers. If teacher A's students outperform teacher B's, is it because method A is better than method B, or because teacher A is more effective than teacher B, or both? There is no way to tell which of these alternative explanations to choose because the effects of teachers and methods are inseparably mixed together---that is, confounded. It is worth noting that, if the students all perform equally well, it may still be true, for example, that method A is better than method B, but teacher B is more talented than teacher A, and the effects have "canceled each other out".

---

Biometrics Unit, Department of Plant Breeding and Biometry, Cornell University, Ithaca, New York.

Any reasonable comparison of procedures must avoid confounding teaching methods with types of students. Similar groups of students should experience both methods, and these groups should be representative of the student population on which the teaching devices may be used. To use the new method on an honor section or class of volunteers would be hopelessly naive if comparison was intended.

A less trivial example of confounding may be seen when the originator of a new method is to implement the method himself, and try to compare it with another teaching procedure. If the innovator teaches by his proposed method while another teacher uses the other approach being considered, then confounding of methods with teachers is present. If the innovator attempts to achieve a reasonable comparison by teaching one or more groups with each method, the relative effectiveness of the methods is confounded with his inability to teach them in such a way as to allow a fair comparison. Because of his emotional investment in the new procedure, the innovating teacher cannot hope to teach both methods with the same enthusiasm and interest. If the new approach appears to succeed, it may not be clear whether the new procedure is intrinsically superior or whether it is simply presented more effectively. If the new procedure is not intended for use by teachers other than the innovator, then perhaps it is successful even if its advantages stem solely from the improved attitude of the teacher. However, if the innovator hopes to see his method used successfully by other teachers, then he must be concerned with the intrinsic superiority of his approach, since other teachers may not share his enthusiasm for the innovation.

It may be that the innovation is to be implemented by teachers other than the innovator, as was the case, for example, with the new high school biology curriculum. In this case one might consider randomly assigning students to classes and teachers to teaching methods. The new method may be greeted by a certain amount of positive or negative enthusiasm on the part of those teachers who must change their teaching habits and adopt the new approach. Comparison between methods will again be confounded with teacher attitudes. However, this confounding is not disastrous as it may add to the reality of the comparison. The new method must in practice withstand this problem of attitudes, at least in its first several years of use.

Finally, one must be cautious about confounding arising through the "Hawthorne Effect". Students being given the new method may sense that they are objects of

special interest and are receiving special attention. Their performance may improve as an artifact of this alone, even when there is truly no difference in the teaching methods.

#### WHEN GOALS HAVE CHANGED

Often a change in teaching practices occurs in conjunction with a redefinition and revision of goals. For example, emphasis may shift from teaching facts to teaching concepts. Since the old and new methods are not designed to achieve the same goal, no comparison of methods is relevant. Instead, one should be evaluating and comparing the alternative goals, and this must be done subjectively.

#### REFLECTIONS

Comparing two procedures, which are designed to achieve the same goal, is relevant but prone to confounding. However, two situations seem to exist where comparison is still possible. If the innovator anticipates that he alone will use his new method, then he may compare procedures by teaching some students with each method and evaluating the degree to which the two groups achieve the goals he has set. If the innovation is to be widely implemented by teachers not involved in its creation, it can be implemented by randomly selected teachers and the performance of the students under the two procedures can be compared. In both of these situations there is confounding of teaching methods with teacher attitudes, but in a way which preserves the relevance of the comparison in terms of estimating the actual usefulness of the procedures in practice.

When one's changing goals have given rise to the innovation, then the old and new method do not share a common purpose. Teaching methods should not be compared unless one has several alternative innovative procedures in mind for reaching the new goal. Instead, a subjective evaluation of the new goals is relevant.